False Data Injection Attacks and the Distributed Countermeasure in DC Microgrids

Mengxiang Liu, Chengcheng Zhao, Ruilong Deng, Peng Cheng, and Jiming Chen

***Abstract*—In this paper, we consider a hierarchical control based DC microgrid (DCmG) equipped with unknown input observer (UIO) based detectors, where the potential false data injection (FDI) attacks and the distributed countermeasure are investigated. First, we find that the vulnerability of the UIO-based detector originates from the lack of knowledge of true inputs. Zero trace stealthy (ZTS) attacks can be launched by secretly faking the unknown inputs, under which the detection residual will not be altered, and the impact on the DCmG in terms of voltage balancing and current sharing is theoretically analyzed. Then, to mitigate the ZTS attack, we propose an automatic and timely countermeasure based on the average point of common coupling (PCC) voltage obtained from the dynamic average consensus (DAC) estimator. The integrity of the communicated data utilized in DAC estimators is guaranteed via UIO-based detectors, where the DAC parameters are perturbed in a fixed period to be concealed from attackers. Finally, the detection and mitigation performance of the proposed countermeasure is rigorously investigated, and extensive simulations are conducted in Simulink/PLECS to validate the theoretical results.**

***Index Terms*—DC Microgrid; False data injection attack; Unknown input observer; Distributed countermeasure.**

I. Introduction

During the past decade, the microgrid composed of distributed generation units (DGUs), storage devices, and flexible loads has become one of the most promising solutions to integrate DGUs such as photovoltaic (PV) panels and wind turbines into the power distribution system [2]. In particular, the tremendous growth in DC loads such as laptop computers, LED lights, and telecommunication centers indicates that the DC microgrid (DCmG) would be an economic and feasible solution in addressing the future energy needs [3].

In DCmGs, the hierarchical control framework is typically adopted to achieve the overall objective such as voltage balancing and current sharing [4]. Specifically, the primary control layer regulates the output voltage of the buck converter to track the reference point of common coupling (PCC) voltage. The secondary control layer adjusts the reference PCC voltage by employing centralized or distributed communication networks [5], to improve the accuracy of current sharing. However, the adoption of information and communications technology also brings in new vulnerabilities like the threats of malicious cyberattacks, which could cause economic losses to or even crash the DCmG. Since there exist many special characteristics unique to the DCmG compared with the general cyber-physical system (CPS) like the high interconnectivity, the hierarchical control framework, the flexible network topology [6], and etc., considerable attention has been attracted to the unique cybersecurity issue therein.

In the power and energy society, the topic of the cybersecurity issue in microgrids has received widespread attention. Considering the microgrid operating in the autonomous mode, Zhang *et al.* [7] investigated the impact of false data injection (FDI) attacks on distributed load sharing and derived the stable regions under attacks. For a well-planned set of *balanced* FDI attacks where no physical error is incurred in the DCmG, Sahoo *et al.* [8] proposed a cooperative vulnerability factor based anomaly detection framework. In [9], Beg *et al.* proposed a *signal temporal logic* based attack detection framework in the DCmG, which can monitor the output voltages and currents against predefined specifications. Zhao *et al.* [10] proposed an adaptive resilient control scheme for the variable-speed wind turbine operating at low-speed region in face of FDI attacks. Nevertheless, the aforementioned literature does not consider the possibility of intelligent attackers, nor investigates the corresponding countermeasure. The intelligent attacker is likely to bypass a certain detector after fully understanding the system model knowledge, and cause specific and accurate adverse effect without being detected. Recent security incidents showed that the intelligent attacker can learn necessary information after penetrating into the system, or collect them from insiders, who have access to critical information legally [11]. Hence, it is of great significance to study the possible threats that could be caused by intelligent attacks, and propose the corresponding countermeasure accordingly.

Since the DCmG is a typically CPS, we also review representative literature about the cybersecurity issue in the context of CPSs. In [12], Pasqualetti *et al.* characterized the undetectable attacks in terms of *zero dynamics*, and designed centralized and distributed attack detection monitors. Inspired by the model-based fault diagnosis technique [13], Teixeira *et al.* proposed a distributed scheme to detect and isolate cyberattacks utilizing the unknown input observer (UIO), which requires that each agent should have certain global knowledge [14], [15]. Nevertheless, the aforementioned methods either rely on the centralized entity or require that each agent should have certain global knowledge, which may be not compatible with the scalability property required by the DCmG [16]. Moreover, Barboni *et al.* [17] designed a novel distributed observer-based estimation technique for detecting covert attacks, and thoroughly investigated the sufficient detectability conditions. Yet, merely local covert attacks inside the subsystem were considered. Recently, Gallo *et al.* [18] proposed a

This work was supported in part by the Science and Technology Innovation 2030 Program under Grant 2018AAA0101605, in part by the National Natural Science Foundation of China under Grants 61833015, 62073285, 62061130220, 61903328, in part by the Zhejiang Provincial Natural Science Foundation under Grants LZ21F020006, LZ22F030010, and in part by the Fundamental Research Funds for the Central Universities (226-2022-00120). (*Corresponding author: Ruilong Deng*)

A preliminary version of this paper was presented at the IEEE American Control Conference, Philadelphia, PA, July, 2019 [1]. The authors are with State Key Lab. of Industrial Control Technology, Zhejiang University, Hangzhou, China (e-mails: {lmx329, chengchengzhao, dengruilong, lunarheart, and cjm}@zju.edu.cn)

completely distributed monitoring scheme by combining the Luenberger observers with UIOs, which solely requires the local model knowledge and local information flow, and can be directly applied to the DCmG for the integrity validation of the communicated data between DGUs. However, it is worth noting that there still exist cyberattacks unforeseeable to the proposed monitoring scheme, and the impact of such attacks has not yet been investigated and mitigated.

Towards this end, in this paper, we investigate the vulnerability of the UIO-based detector, and theoretically analyze the threat of such vulnerability in the context of DCmGs. Furthermore, based on the analysis, we propose an automatic and timely countermeasure against the vulnerability, and the performance in vulnerability perception and threat mitigation is thoroughly studied. In addition to our preliminary work [1], we design a distributed countermeasure and provide rigorous proofs for the theoretical results. Specifically, the contributions of this paper are listed as follows:

1) We find that the vulnerability of the UIO-based detector originates from the lack of knowledge of true inputs. By secretly faking the unknown inputs, the zero trace stealthy (ZTS) attack can be launched without altering the detection residual. Moreover, we theoretically analyze the impact of both single and cooperative ZTS attacks on the DCmG.
2) Based on the average PCC voltage (APV) obtained from the dynamic average consensus (DAC) estimator, we propose an automatic and timely countermeasure against ZTS attacks. The DAC parameters are perturbed in a fixed period to be concealed from the attacker, such that the integrity of the communicated data utilized in DAC estimators can be guaranteed via UIO-based detectors.
3) The sufficient condition on detecting ZTS attacks is derived, and the effectiveness of the impact mitigation strategy is rigorously analyzed. Extensive simulations are conducted in Matlab Simulink/PLECS to validate the theoretical results.

The rest of this paper is organized as follows. Section II presents the system model and the problem formulation. Section III illustrates the construction of ZTS attacks, and investigates the impact of ZTS attacks on DCmGs. The distributed countermeasure is proposed and elaborated in Section IV. Finally, simulations results are shown in Section V and Section VI concludes this paper.

Notation: $\mathbb{C}$ is the set of complex numbers, and $\mathbb{R}/\mathbb{R}^n$ is the set of real numbers/vectors. The symbol $|\cdot|$ denotes the cardinality of a finite set and component-by-component absolute value of a matrix/vector, and $\|\cdot\|$ represents the norm of a matrix/vector. Inequalities of matrices/vectors are compared component-by-component, and $\lim_{t\to\infty} y(t)$ is denoted by $y(\infty)$ for brevity. Let $\mathbb{1}^n/\mathbb{1}^{n\times n}$ and $\mathbb{0}^n/\mathbb{0}^{n\times n}$ denote vectors/matrices with all 1 and 0 entries, respectively, and I^n denotes the unit matrix with $n\times n$ dimension. Scalar $v_{[m]}$ denotes the m-th entry of vector $\boldsymbol{v}\in\mathbb{R}^n$. Let $\mathbb{H}^1$ denote the subspace of $\mathbb{R}^n$ composed by all vectors satisfying $\langle\boldsymbol{v}\rangle=\frac{1}{n}\sum_{i=1}^n v_{[i]}=0$, where $\langle\boldsymbol{v}\rangle$ denotes the average of elements in vector $\boldsymbol{v}$.

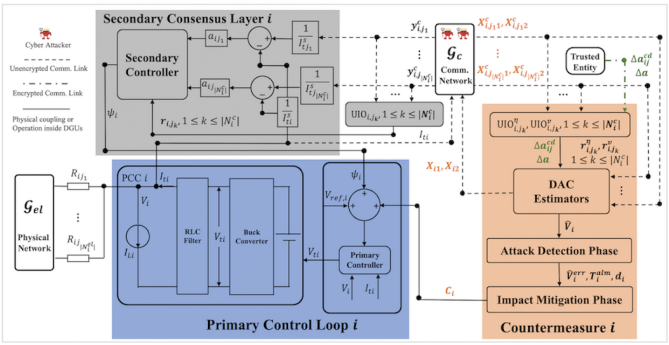


Fig. 1. This figure shows the hierarchical control framework and the distributed countermeasure in DGU i.

Intuitively, each vector in $\mathbb{H}^1$ has $n-1$ freedom[1], indicating that the dimension of $\mathbb{H}^1$ is $n-1$, i.e., $\dim\{\mathbb{H}^1\}=n-1$. Moreover, let $\mathbb{H}^1_\perp$ be the orthogonal subspace of $\mathbb{H}^1$ such that $\mathbb{H}^1\bigoplus\mathbb{H}^1_\perp=\mathbb{R}^n$, then we have $\forall\boldsymbol{v}\in\mathbb{H}^1_\perp, \boldsymbol{v}=\alpha\mathbb{1}^n, \alpha\in\mathbb{R}$ and $\dim\{\mathbb{H}^1_\perp\}=1$. Here $\bigoplus$ denotes the direct sum of subspaces.

II. System Model and Problem Formulation

A. Network Model

Graph Theory: A weighted undirected graph (WUG) is denoted by $\mathcal{G}=\{\mathcal{V},\mathcal{E},W\}$, where $\mathcal{V}=\{1,2,\cdots,n\}$ is the set of nodes, $\mathcal{E}=\{(i,j)\}\subseteq\mathcal{V}\times\mathcal{V}$ is the set of edges, and $W=\mathrm{diag}\{a_{ij}\}\in\mathbb{R}^{|\mathcal{E}|\times|\mathcal{E}|}$ is the diagonal matrix composed by edge weights $a_{ij},\forall(i,j)\in\mathcal{E}$. The set of neighbors of node i is $\mathcal{N}_i=\{j|(i,j)\in\mathcal{E}\}$. After assigning each edge of $\mathcal{G}$ an arbitrary direction, the oriented incidence matrix is computed as $P(\mathcal{G})\in\mathbb{R}^{|\mathcal{V}|\times|\mathcal{E}|}$, where the order of columns corresponds to the order of edge weights in W [19]. Then, the Laplacian matrix of $\mathcal{G}$ can be expressed as $\mathscr{L}(\mathcal{G})=P(\mathcal{G})WP(\mathcal{G})^{\mathrm{T}}$, which is independent of the edge orientations.

Electrical and Communication Networks: The electrical network of DCmG is represented by WUG $\mathcal{G}_{el}=\{\mathcal{V},\mathcal{E}_{el},W_{el}\}$, where nodes are DGUs, edges are power lines whose orientations define reference directions for positive currents. Moreover, edge weights are conductances of power lines. The set of neighbors of node i is $\mathcal{N}_i^{el}$, $|\mathcal{V}|=N$ and the Laplacian matrix is $\mathscr{L}(\mathcal{G}_{el})=M$. The communication network of DCmG is denoted by WUG $\mathcal{G}_c=\{\mathcal{V},\mathcal{E}_c,W_c\}$, where edges are communication links and edge weights are $a_{ij}^c,\forall(i,j)\in\mathcal{E}_c$. The set of neighbors of node i is $\mathcal{N}_i^c$, and the Laplacian matrix is $\mathscr{L}(\mathcal{G}_c)=L$.

B. DGU Dynamics

As shown in Fig. 1, each DGU contains a DC voltage source, a buck converter, a local load current, and a RLC (resistor, inductor, and capacitor) filter. Notice that V_{ti} is the output voltage of buck converter and I_{Li} is the constant load current. Moreover, V_i and I_{ti} are the PCC voltage and the output current, respectively. The hierarchical control framework is deployed in each DGU, where the primary controller tracks the local reference PCC voltage and the secondary consensus layer regulates the local reference PCC voltage to achieve

[1]Any $n-1$ entries in the vector can be set arbitrarily.

current sharing and voltage balancing in the DCmG [6]. The dynamical model of DGU $i \in \mathscr{V}$ is

$$\begin{cases} \dot{\boldsymbol{x}}_i(t) = A_{ii}\boldsymbol{x}_i(t) + \boldsymbol{b}_i u_i(t) + \boldsymbol{g}_i\psi_i(t) + \\ \qquad + M_i\boldsymbol{d}_i + \boldsymbol{\xi}_i(t) + \boldsymbol{\omega}_i(t), \\ \boldsymbol{y}_i(t) = \boldsymbol{x}_i(t) + \boldsymbol{\rho}_i(t), \end{cases} \tag{1}$$

where $\boldsymbol{x}_i(t) = [V_i(t), I_{ti}(t), v_i(t)]^{\mathrm{T}}$ is the state vector, and $v_i(t)$ is the integral of the voltage tracking error defined by $\dot{v}_i(t) = V_{ref,i} + \psi_i(t) - V_i(t)$. Here $V_{ref,i}$ is the nominal reference PCC voltage and $\psi_i(t)$ is the secondary control input. Moreover, $\boldsymbol{d}_i = [I_{Li}, V_{ref,i}]^{\mathrm{T}}$ is the constant exogenous input vector, and $\boldsymbol{y}_i(t) \in \mathbb{R}^3$ is the output vector. The physical couplings with neighboring DGUs are modeled as $\boldsymbol{\xi}_i(t) = \sum_{j\in\mathscr{N}_i^{el}} A_{ij}\boldsymbol{x}_j(t) \in \mathbb{R}^3$. The primary control input is computed as

$$u_i(t) = V_{ti} = \boldsymbol{k}_i^{\mathrm{T}}\boldsymbol{y}_i(t), \tag{2}$$

where the primary control gain $\boldsymbol{k}_i \in \mathbb{R}^3$ depends merely on the model knowledge of DGU i and the interconnected power lines [16]. The secondary control input is obtained through the following consensus scheme, i.e.,

$$\dot{\psi}_i(t) = -[0, k_I, 0] \sum_{j\in\mathscr{N}_i^c} a_{ij}^c \left(\frac{\boldsymbol{y}_i(t)}{I_{ti}^s} - \frac{\boldsymbol{y}_{i,j}^c(t)}{I_{tj}^s}\right), \tag{3}$$

where $\boldsymbol{y}_{i,j}^c(t)$ is the output of DGU j communicated to DGU i, $I_{ti}^s > 0$ and $I_{tj}^s > 0$ are rated currents corresponding to DGU i and DGU j, respectively, and $k_I > 0$ is the weight parameter invariant among all DGUs. We have the following Assumptions regarding to the DCmG model.

Assumption 1: The process noise and measurement noise are unknown-but-bounded i.e., $|\boldsymbol{\omega}_i(t)| \le \bar{\boldsymbol{\omega}}_i \in \mathbb{R}^3, |\boldsymbol{\rho}_i(t)| \le \bar{\boldsymbol{\rho}}_i \in \mathbb{R}^3, \forall t \ge 0$.

Assumption 2: The nominal reference PCC voltages are equal among all DGUs, i.e., $V_{ref,i} = V_{ref}, \forall i \in \mathscr{V}$.

Assumption 3: The WUGs $\mathscr{G}_c$ and $\mathscr{G}_{el}$ are both connected, and they have the same topology and edge weights ($L = M$).

Under Assumptions 1-3, the hierarchical control framework (2)-(3) can achieve voltage balancing and current sharing in DCmGs [6], which are formally defined below.

Definition 1 (Voltage Balancing): Under Assumption 2, voltage balancing is achieved if $\langle \boldsymbol{v}(\infty) \rangle = V_{ref}$, where $\boldsymbol{v}(t) = [V_i(t), \cdots, V_N(t)]^{\mathrm{T}}$, and $\langle \boldsymbol{v}(\infty) \rangle$ denotes the steady-state APV.

Definition 2 (Current Sharing): For constant load currents, current sharing is achieved if $\frac{I_{ti}(\infty)}{I_{ti}^s} = \frac{I_{tj}(\infty)}{I_{tj}^s}, \forall i, j \in \mathscr{V}$, i.e., load currents are shared proportionally to the rated currents.

C. Attack Model

In this paper, we consider FDI attacks injecting malicious signals into communication links between DGUs. In particular, the FDI attack against $(i, j) \in \mathscr{E}_c$ is modeled as

$$\boldsymbol{y}_{i,j}^c(t) = \boldsymbol{y}_j(t) + \beta(t - T_a)\boldsymbol{\phi}_{i,j}(t), \tag{4}$$

where $\boldsymbol{\phi}_{i,j}(t)$ is an arbitrary vector designed by the attacker, and $\beta(t - T_a)$ is the step function with T_a delay. The attack is started at $t = T_a$, i.e., $\boldsymbol{y}_{i,j}^c(t) = \boldsymbol{y}_j(t), \forall t \le T_a$. In this study, we consider the continuous and differentiable attack vector and give the following assumption.

Assumption 4: The attack vector $\boldsymbol{\phi}_{i,j}(t)$ is continuous and differentiable.

Remark 1: Assumption 4 is practical as it can guarantee the smoothness of the corrupted signal, such that the corrupted signal would be indistinguishable from the normal signal. Moreover, the resulting conclusion under Assumption 4 would have explicit forms and could further facilitate our future research on more general attack vectors.

Moreover, the attacker is likely to obtain system parameters from the insider [20], who can get access to them legally. But the attacker is hard to obtain real-time system parameters as the insider only discloses them to the attacker in a specific time period to guarantee his/her hiddenness. Therefore, we consider that the attacker is able to obtain system parameters involved in DGU dynamics (1) every few hours or days (not immediately). Moreover, the attacker is able to eavesdrop the communicated data through IP spoofing attacks. Nevertheless, the attacker cannot intrude into DGU i or the DCmG control center[2] due to various host-based defense mechanisms [22], indicating that the primary control input $u_i(t)$ will not be compromised.

D. UIO-based Detector

According to [18], a bank of UIOs are deployed in each DGU to identify and isolate the FDI attacks among the neighboring communication links. For convenience, the dynamical model of DGU $j \in \mathscr{N}_i^c$ (1) is transformed to

$$\begin{cases} \dot{\boldsymbol{x}}_j(t) = A_{kj}\boldsymbol{x}_j(t) + \bar{E}_j\bar{\boldsymbol{d}}_j(t) + \boldsymbol{\omega}_j(t) + \boldsymbol{b}_j\boldsymbol{k}_j^{\mathrm{T}}\boldsymbol{\rho}_j(t), \\ \boldsymbol{y}_j(t) = \boldsymbol{x}_j(t) + \boldsymbol{\rho}_j(t), \end{cases} \tag{5}$$

where $A_{kj} = A_{jj} + \boldsymbol{b}_j\boldsymbol{k}_j^{\mathrm{T}} \in \mathbb{R}^{3\times3}$, $\bar{E}_j\bar{\boldsymbol{d}}_j(t) = M_j\boldsymbol{d}_j + \boldsymbol{g}_j\psi_j(t) + \boldsymbol{\xi}_j(t)$, and $\bar{E}_j \in \mathbb{R}^{3\times2}$ is a full column rank matrix related to the capacitor parameter C_{tj} as shown in (6).

$$\bar{E}_j = \begin{bmatrix} \frac{1}{C_{tj}} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{\mathrm{T}}, H_j = \begin{bmatrix} 1 & 0 & 0 \\ h_{12} & h_{22} & h_{32} \\ 0 & 0 & 1 \end{bmatrix}^{\mathrm{T}}. \tag{6}$$

Moreover, vector $\bar{\boldsymbol{d}}_j(t)$ contains the inputs of DGU j unknown to DGU i. Based on (5), one can easily verify that $\mathrm{rank}(\mathrm{I}^3\bar{E}_j) = \mathrm{rank}(\bar{E}_j)$ and matrix $\begin{bmatrix} s\mathrm{I}^3 - A_{kj} & \bar{E}_j \\ \mathrm{I}^3 & \mathbb{0}^{3\times2} \end{bmatrix}$ has full column rank $\forall s \in \mathbb{C}$. Hence, according to Theorem 1 in [23], the full order UIO in DGU i can be constructed as

$$\mathrm{UIO}_{i,j} \begin{cases} \dot{\boldsymbol{z}}_{i,j}(t) = F_j\boldsymbol{z}_{i,j}(t) + \hat{K}_j\boldsymbol{y}_{i,j}^c(t), \\ \hat{\boldsymbol{x}}_{i,j}(t) = \boldsymbol{z}_{i,j}(t) + H_j\boldsymbol{y}_{i,j}^c(t), \end{cases} \tag{7}$$

under which, in the normal case, the estimated state $\hat{\boldsymbol{x}}_{i,j}(t)$ will converge asymptotically to $\boldsymbol{x}_j(t)$ regardless of the unknown input vector $\bar{\boldsymbol{d}}_j(t)$. Here $\boldsymbol{z}_{i,j}(t) \in \mathbb{R}^3$ is the internal

[2]In the DCmG, the control center is mainly responsible for the tertiary control layer including optimal operation in grid-tied and islanded operating modes and power flow control in grid-tied mode [21].

state of UIO (7), and UIO parameters $F_j, \hat{K}_j, H_j \in \mathbb{R}^{3\times 3}$ need to satisfy

$$T_j\bar{E}_j = \mathbb{0}^{3\times 2}, \tag{8a}$$
$$T_j = \mathrm{I}^3 - H_j, \tag{8b}$$
$$\hat{K}_j = K_{j1} + K_{j2}, \tag{8c}$$
$$F_j = T_jA_{kj} - K_{j1}, \tag{8d}$$
$$K_{j2} = F_jH_j, \tag{8e}$$

where $K_{j1}, K_{j2}, T_j \in \mathbb{R}^{3\times 3}$, H_j is defined in (6), h_{12}, h_{22}, h_{32} are arbitrary scalars, and K_{j1} should be appropriately chosen to make the eigenvalues of F_j all lie in the open left half-plane based on (8d). In the absence of attacks, the analytical expression of detection residual $\boldsymbol{r}_{i,j}(t) = \boldsymbol{y}^c_{i,j}(t) - \hat{\boldsymbol{x}}_{i,j}(t)$ can be obtained given DGU dynamics (5) and UIO (7), i.e.,

$$\boldsymbol{r}_{i,j}(t) = e^{F_jt}\big(\boldsymbol{\sigma}_{2i,j}(0) + \boldsymbol{\sigma}_{3i,j}(t)\big) + T_j\boldsymbol{\rho}_j(t), \tag{9}$$

where $\boldsymbol{\sigma}_{2i,j}(0) = \boldsymbol{x}_j(0) - \hat{\boldsymbol{x}}_{i,j}(0) + H_j\boldsymbol{\rho}_j(0)$ and $\boldsymbol{\sigma}_{3i,j}(t) = \int_0^t e^{-F_j\tau}\big(T_j\boldsymbol{\omega}_j(\tau) + (T_j\boldsymbol{b}_j\boldsymbol{k}_j - \hat{K}_j)\boldsymbol{\rho}_j(\tau)\big)\mathrm{d}\tau$. Since $\boldsymbol{y}_j(0) = \boldsymbol{x}_j(0) + \boldsymbol{\rho}_j(0)$, $\boldsymbol{z}_{i,j}(0)$ is set as $T_j\boldsymbol{y}^c_{i,j}(0)$ such that the initial state estimation error can be bounded by the bound of measurement noise in the absence of attacks, i.e.,

$$|\boldsymbol{x}_j(0) - \hat{\boldsymbol{x}}_{i,j}(0)| = |\boldsymbol{y}_j(0) - \boldsymbol{y}^c_{i,j}(0) - \boldsymbol{\rho}_j(0)| = |\boldsymbol{\rho}_j(0)| \le \bar{\boldsymbol{\rho}}_j.$$

Moreover, as F_j is Hurwitz stable, there exist positive constants κ, μ such that $||e^{F_jt}|| \le \kappa e^{-\mu t}, \forall t \ge 0$. Then, the time-varying detection threshold $\bar{\boldsymbol{r}}_{i,j}(t)$ is computed such that

$$|\boldsymbol{r}_{i,j}(t)| \le \bar{\boldsymbol{r}}_{i,j}(t) = \kappa e^{-\mu t}\big(\bar{\boldsymbol{\sigma}}_{2i,j}(0) + \bar{\boldsymbol{\sigma}}_{3i,j}(t)\big) + |T_j|\bar{\boldsymbol{\rho}}_j \tag{10}$$

always hold in the absence of attacks, where $|\boldsymbol{\sigma}_{2i,j}(0)| \le \bar{\boldsymbol{\sigma}}_{2i,j}(0) = (\mathrm{I}^3 + |H_j|)\bar{\boldsymbol{\rho}}_j$ and $|\boldsymbol{\sigma}_{3i,j}(t)| \le \bar{\boldsymbol{\sigma}}_{3i,j}(t) = \int_0^t |e^{-F_j\tau}|(|T_j|\bar{\boldsymbol{\omega}}_j + |T_j\boldsymbol{b}_j\boldsymbol{k}_j - \hat{K}_j|\bar{\boldsymbol{\rho}}_j)\mathrm{d}\tau$.

Once (10) is violated, it is considered that the data $\boldsymbol{y}^c_{i,j}(t)$ received from DGU j is corrupted by attacks. With some abuse of the notation, let $\boldsymbol{r}_{i,j}(t)$ be the detection residual under attacks, and it is decomposed as

$$\boldsymbol{r}_{i,j}(t) = \tilde{\boldsymbol{r}}_{i,j}(t) + \boldsymbol{r}^a_{i,j}(t),$$

where $\tilde{\boldsymbol{r}}_{i,j}(t)$ is the healthy residual component equating to (9) and $\boldsymbol{r}^a_{i,j}(t)$ is the malicious component associated with attacks. Given the attack model (4), DGU dynamics (5) and UIO (7), we obtain

$$\begin{aligned}\boldsymbol{r}^a_{i,j}(t) = e^{F_j(t-T_a)}H_j\boldsymbol{\phi}_{i,j}(T_a) + T_j\boldsymbol{\phi}_{i,j}(t) + \\ - \int_{T_a}^t e^{F_j(t-\tau)}\hat{K}_j\boldsymbol{\phi}_{i,j}(\tau)\mathrm{d}\tau.\end{aligned} \tag{11}$$

E. Problems of Interest

In this paper, we are interested in the FDI attacks that cause no impact on the detection residual, i.e., $\boldsymbol{r}^a_{i,j}(t) = \mathbb{0}^3$, while the received output $\boldsymbol{y}^c_{i,j}(t)$ deviates a lot from the true one $\boldsymbol{y}_j(t)$. For clarity, we define the FDI attack aforementioned as

Definition 3 (ZTS Attack): Given DGU dynamics (5) and UIO-based detector (7), the FDI attack (4) is **ZTS** if

$$\begin{cases}\boldsymbol{\phi}_{i,j}(t) \ne \mathbb{0}^3, \exists t \ge 0, \\ \boldsymbol{r}^a_{i,j}(t) = \mathbb{0}^3, \forall t \ge 0.\end{cases}$$

Remark 2: Zero-dynamics attacks characterize a class of undetectable attacks that excite only *zero dynamics* of a dynamical system, which can make the system states diverge while leaving no trace on the outputs, and thus are inherently undetectable for detectors. According to the attack model in [12], the ZTS attack corrupting the outputs of a dynamical system can be described by the Rosenbrock matrix $P(s) = \begin{bmatrix} s\mathrm{I}^3 - A_{kj} & \mathbb{0}^{3\times 3} \\ \mathrm{I}^3 & \mathrm{I}^3 \end{bmatrix}$, under which the *zero dynamics* of the system can never be excited with $B = \mathbb{0}^{3\times 3}$. Hence, ZTS attacks are essentially different from zero-dynamics attacks. Specifically, ZTS attacks reveal the vulnerability of the attack detection and identification for a dynamical system when there exist some *unknown inputs* regardless of the *zero dynamics*.

The following three problems are formulated: (1) How can the attacker construct ZTS attacks? **(P1)** (2) How will ZTS attacks affect the DCmG? **(P2)** (3) How to detect and mitigate the impact of ZTS attacks? **(P3)**

III. ZTS Attack and the Impact Analysis

In this section, we characterize the condition for the FDI attack (4) to be ZTS and investigate the impact of ZTS attacks on DCmGs.

A. ZTS Attack

Although DGU i can estimate the unknown inputs of DGU $j \in \mathcal{N}_i$ from $\boldsymbol{y}^c_{i,j}(t)$, it is still not sure whether the estimated unknown inputs are *true* or not. Hence, the intuition is to deceive DGU i utilizing a fake unknown input vector $\bar{\boldsymbol{d}}^a_{ij}(t)$, which motivates the following analysis.

Theorem 1: Given DGU dynamics (5) and UIO-based detector (7), the FDI attack (4) under Assumption 4 is ZTS if and only if the attack vector $\boldsymbol{\phi}_{i,j}(t), \forall t \ge T_a$ satisfies

$$\begin{cases}\dot{\boldsymbol{\phi}}_{i,j}(t) = A_{kj}\boldsymbol{\phi}_{i,j}(t) + \bar{E}_j\bar{\boldsymbol{d}}^a_{ij}(t), \\ \boldsymbol{\phi}_{i,j}(T_a) = \mathbb{0}^3,\end{cases} \tag{12}$$

where $\bar{\boldsymbol{d}}^a_{ij}(t)$ should satisfy $\bar{E}_j\bar{\boldsymbol{d}}^a_{ij}(t) \ne \mathbb{0}^3$.

Proof: *(If)* According to the DGU dynamics (5), the Laplace form of $\boldsymbol{y}^c_{i,j}(t)$ corrupted by the attack vector $\boldsymbol{\phi}_{i,j}(t)$ satisfying (12) is

$$\begin{aligned}\boldsymbol{y}^c_{i,j}(s) = (s\mathrm{I}^3 - A_{kj})^{-1}(\boldsymbol{x}_j(0) + \bar{E}_j\tilde{\boldsymbol{d}}_j(s) + \\ + \boldsymbol{\omega}_j(s) + \boldsymbol{b}_j\boldsymbol{k}_j\boldsymbol{\rho}_j(s)) + \boldsymbol{\rho}_j(s),\end{aligned} \tag{13}$$

where $\tilde{\boldsymbol{d}}_j(s) = \bar{\boldsymbol{d}}_j(s) + \bar{\boldsymbol{d}}^a_{ij}(s)$ integrates the normal and fake unknown input vectors. It follows from (13) that $\boldsymbol{y}^c_{i,j}(t)$ can be interpreted as the output of system (5) whose unknown input vector $\bar{\boldsymbol{d}}_j(t)$ is switched to $\tilde{\boldsymbol{d}}_j(t)$ at $t = T_a$. Hence, the attack vector $\boldsymbol{\phi}_{i,j}(t)$ satisfying (12) will not alter the detection residual $\boldsymbol{r}_{i,j}(t)$, which is designed to be insensitive to the variation of unknown inputs. The proof of the sufficient part is completed.

(Only If) Suppose that there exists the FDI attack (4) such that $\boldsymbol{r}^a_{i,j}(t) = \mathbb{0}^3, \forall t \ge T_a$. Substituting (11) into $\boldsymbol{r}^a_{i,j}(T_a) = \mathbb{0}^3$, we obtain $(H_j + T_j)\boldsymbol{\phi}_{i,j}(T_a) = \mathbb{0}^3$, under which we can

derive $\boldsymbol{\phi}_{i,j}(T_a) = \mathbb{0}^3$ based on (8b). Then, integrating (11) with $\boldsymbol{r}^a_{i,j}(t) = \mathbb{0}^3$, we have

$$T_j\boldsymbol{\phi}_{i,j}(t) = \int_{T_a}^{t} e^{F_j(t-\tau)}\hat{K}_j\boldsymbol{\phi}_{i,j}(\tau)\mathrm{d}\tau. \tag{14}$$

Given Assumption 4, (14) can be achieved only if $\boldsymbol{\phi}_{i,j}(t)$ satisfies

$$T_j\dot{\boldsymbol{\phi}}_{i,j}(t) = (F_jT_j + \hat{K}_j)\boldsymbol{\phi}_{i,j}(t), \tag{15}$$

which is obtained by calculating the differentials of both sides of equation (14). It is noted that the vector $\boldsymbol{\phi}_{i,j}(t)$ belongs to the null space of composite matrix $[\hat{K}_j; T_j]$ must satisfy (15) as $T_j\boldsymbol{\phi}_{i,j}(t) = \hat{K}_j\boldsymbol{\phi}_{i,j}(t) = \mathbb{0}^3$ can immediately establish (15). Based on equations (8b)-(8e), we have

$$F_jT_j + \hat{K}_j = F_j(\mathrm{I}^3 - H_j) + \hat{K}_j = F_j + K_{j1} = T_jA_{kj}. \tag{16}$$

Substituting (16) into (15), we obtain

$$T_j(\dot{\boldsymbol{\phi}}_{i,j}(t) - A_{kj}\boldsymbol{\phi}_{i,j}(t)) = \mathbb{0}^3, \forall t \geq T_a. \tag{17}$$

It follows from equations (6), and (8a)-(8b) that $T_j\bar{E}_j = \mathbb{0}^{3\times 2}$ and $\mathrm{rank}(T_j) + \mathrm{rank}(\bar{E}_j) = 3$, indicating that the null space of T_j coincides with the range space of $\bar{E}_j$. Thus, (17) is equivalent to

$$\dot{\boldsymbol{\phi}}_{i,j}(t) = A_{kj}\boldsymbol{\phi}_{i,j}(t) + \bar{E}_j\bar{\boldsymbol{d}}^a_{ij}(t), \forall t \geq T_a, \tag{18}$$

implying that $\boldsymbol{\phi}_{i,j}(t)$ will not be zero persistently once $\bar{E}_j\bar{\boldsymbol{d}}^a_{ij}(t) \neq \mathbb{0}^3$. The proof of the necessary part is completed.■

Based on (12), the attacker can construct ZTS attacks once he/she could get access to $A_{kj}, \bar{E}_j$, which are determined by electrical parameters (resistance, capacitance, and inductance) and the primary control gain ($\boldsymbol{k}_j$).

Remark 3: Under Assumption 4, ZTS attacks can only be constructed by utilizing the fake unknown input $\boldsymbol{d}^a_{ij}(t)$, indicating that the vulnerability of the UIO-based detector (7) originates from the lack of knowledge of true inputs.

Remark 4: Indeed, the ZTS attack is a special case of the covert attack described in [18], with the attack vector being initialized at zero. Under the specific initial condition, we have the sufficient and necessary condition for the attack satisfying Assumption 4 to be ZTS. While the derivation of the sufficient and necessary condition for the covert attack is difficult due to the uncertainty of the initial condition introduced by measurement noises. Moreover, we note that there exist some ZTS attacks whose attack vectors are either discontinuous or non-differentiable, but the investigation on them is still challenging due to the implicit and non-unified attack vector forms and is left as our future work.

B. Attack Impact Analysis

In this section, we theoretically analyze the impact of ZTS attacks on voltage balancing and current sharing. According to (12), the pair $(A_{kj}, \bar{E}_j)$ is controllable, and thus the ZTS attack vector $\boldsymbol{\phi}_{i,j}(t)$ can be arbitrarily large with well-designed $\bar{\boldsymbol{d}}^a_{ij}(t)$. However, $\boldsymbol{\phi}_{i,j}(t)$ should be bounded to make the corrupted measurement $\boldsymbol{y}^c_{i,j}(t)$ physically reachable, given the maximal/minimal PCC voltage and output current for DGU j. Hence, we provide the assumption below.

Assumption 5: The fake unknown input vector involved in (12), i.e., $\boldsymbol{d}^a_{ij}(t)$, is a bounded constant vector.

Remark 5: Since A_{kj} is Hurwitz stable, the attack vector generated by (12) with constant $\boldsymbol{d}^a_{ij}(t)$ will eventually converge. Hence, from the perspective of the attacker, it is practical and useful to set $\boldsymbol{d}^a_{ij}(t)$ as a constant vector, under which a bounded $\boldsymbol{\phi}_{i,j}(t)$ could be generated at his/her will. Moreover, the impact of the ZTS attacks with time-varying $\boldsymbol{d}^a_{ij}(t)$ can be analyzed in a similar way referring to the following results.

In the remainder of this paper, $\bar{\boldsymbol{d}}^a_{ij}$ is utilized to denote the constant vector $\bar{\boldsymbol{d}}^a_{ij}(t)$. With some abuse of notations, let $\boldsymbol{\psi}(t) = [\psi_1(t), \cdots, \psi_N(t)]$ be the secondary control input vector under attacks. Similar to $\boldsymbol{r}_{i,j}(t)$, $\boldsymbol{\psi}(t)$ is decomposed as $\boldsymbol{\psi}(t) = \tilde{\boldsymbol{\psi}}(t) + \boldsymbol{\psi}_a(t)$, where $\tilde{\boldsymbol{\psi}}(t)$ denotes the healthy component and $\boldsymbol{\psi}_a(t)$ is the malicious component associated with attacks.

Theorem 2: Under Assumptions 2-5, any single ZTS attack (12) will cause

$$\langle\boldsymbol{\psi}_a(\infty)\rangle = -\frac{k_Ia^c_{ij}}{NI^s_{tj}}\boldsymbol{k}^{\mathrm{T}}A^{-1}_{kj}(\bar{E}_j\bar{\boldsymbol{d}}^a_{ij}(t-T_a) + A^{-1}_{kj}\bar{E}_j\bar{\boldsymbol{d}}^a_{ij}), \tag{19}$$

where $\boldsymbol{k}^{\mathrm{T}} = [0, 1, 0]$. Intuitively, with nonzero $\boldsymbol{k}^{\mathrm{T}}A^{-1}_{kj}\bar{E}_j\bar{\boldsymbol{d}}^a_{ij}$, neither voltage balancing nor current sharing can be achieved in DCmGs.

Proof: The proof is given in Appendix A. ■

Next, we consider the case where ZTS attacks (12) are injected into multi communication links $\tilde{\mathscr{E}}_c \subseteq \mathscr{E}_c$ cooperatively such that

$$\sum_{(i,j)\in\tilde{\mathscr{E}}_c}\frac{k_Ia^c_{ij}}{I^s_{tj}}\boldsymbol{k}^{\mathrm{T}}A^{-1}_{kj}\bar{E}_j\bar{\boldsymbol{d}}^a_{ij} = 0. \tag{20}$$

Theorem 3: Under Assumptions 2-5, the cooperative ZTS attacks (12) satisfying (20) will cause

$$\langle\boldsymbol{\psi}_a(\infty)\rangle = -\sum_{(i,j)\in\tilde{\mathscr{E}}_c}\frac{k_Ia^c_{ij}}{NI^s_{tj}}\boldsymbol{k}^{\mathrm{T}}A^{-2}_{kj}\bar{E}_j\bar{\boldsymbol{d}}^a_{ij}, \tag{21}$$

under which voltage balancing cannot be achieved if $\sum_{(i,j)\in\tilde{\mathscr{E}}_c}\frac{k_Ia^c_{ij}}{I^s_{tj}}\boldsymbol{k}^{\mathrm{T}}A^{-2}_{kj}\bar{E}_j\bar{\boldsymbol{d}}^a_{ij} \neq 0$ and current sharing cannot be achieved if $\sum_{(i,j)\in\tilde{\mathscr{E}}_c}\frac{k_Ia^c_{ij}}{I^s_{tj}}\boldsymbol{k}^{\mathrm{T}}A^{-1}_{kj}\bar{E}_j\bar{\boldsymbol{d}}^a_{ij}\boldsymbol{l}_i \neq \mathbb{0}^N$.

Proof: The proof is given in Appendix B. ■

Remark 6: From the perspective of the attacker, he/she can choose appropriate attack vectors $\boldsymbol{\phi}_{i,j}(t), \forall(i,j) \in \tilde{\mathscr{E}}_c$ referring to the theoretical results in Theorems 2-3 to achieve his/her malicious goals. Specifically, if the attacker can get access to any communication link $(i,j) \in \mathscr{E}_c$, then the single ZTS attack (12) with nonzero $\boldsymbol{k}^{\mathrm{T}}A^{-1}_{kj}\bar{E}_j\bar{\boldsymbol{d}}^a_{ij}$ can destabilize the DCmG. Moreover, if the attacker can get access to multi communication links $\tilde{\mathscr{E}}_c$ simultaneously, then he/she can launch the cooperative ZTS attacks (12) satisfying (20) to induce accurate and specific adverse impact on voltages and currents.

IV. The Distributed Countermeasure

In this section, we propose an automatic and timely countermeasure against ZTS attacks based on the APV obtained from the DAC estimator. As shown in Algorithm 1, the countermeasure is composed of two phases, i.e., attack detection and impact mitigation. In particular, the former phase is to reveal the existence of ZTS attacks by utilizing the detection indicator derived from the APV. Once the detection indicator exceeds a predefined threshold, the latter phase is activated for compensation until voltage balancing is recovered. In the following subsections, we will introduce the DAC estimator, the attack detection phase, and the impact mitigation phase.

Algorithm 1 The Distributed Countermeasure in DGU $i \in \mathscr{V}$

Input: The PCC voltage $V_i(t)$
Deploy the DAC estimator (22) satisfying (25);
Output: ***Attack Detection Phase***
Calculate the detection indicator $\mathscr{A}_i(t)$ according to (30);
Set the detection threshold $\bar{\mathscr{A}}_i$ according to (31);
if $\mathscr{A}_i(t) > \bar{\mathscr{A}}_i$ **then**
    Activate the impact mitigation phase;
else
    Repeat the detection phase;
end if
Output: ***Impact Mitigation Phase***
Compute the compensation value $C_i(t)$ according to (34);
Add $C_i(t)$ to the secondary control input $\psi_i(t)$ according to (35);
if Condition (37) is satisfied **then**
    ▷ *Judge whether voltage balancing is recovered*
    Goto the detection phase;
else
    Repeat the impact mitigation phase;
end if

A. The DAC Estimator

In this subsection, we introduce the DAC estimator equipped with UIO-based detectors, which are utilized to validate the integrity of the communicated data for the estimator. From [24], the dynamics of the DAC estimator in DGU $i \in \mathscr{V}$ are

$$\begin{cases} \dot{\boldsymbol{X}}_{i1}(t) = A_1 \boldsymbol{X}_{i1}(t) + B_1 (V_i(t) - \gamma \sum\limits_{j \in \mathscr{N}_i^c} a_{ij}^{cd} (\eta_i(t) - \eta_{i,j}^c(t))), \\ \widehat{V}_i(t) = C_1 \boldsymbol{X}_{i1}(t), \end{cases} \tag{22a}$$

and

$$\begin{cases} \dot{\boldsymbol{X}}_{i2}(t) = A_2 \boldsymbol{X}_{i2}(t) + B_2 (\gamma \sum\limits_{j \in \mathscr{N}_i^c} a_{ij}^{cd} (\widehat{V}_i(t) - \widehat{V}_{i,j}^c(t))), \\ \eta_i(t) = C_2 \boldsymbol{X}_{i2}(t), \end{cases} \tag{22b}$$

where $V_i(t)$ is the input signal, $\widehat{V}_i(t) \in \mathbb{R}$ is the estimated APV (output signal), and $\boldsymbol{X}_{i1}(t) \in \mathbb{R}^{n_1}, \boldsymbol{X}_{i2}(t) \in \mathbb{R}^{n_2}$ are the internal states. Here n_1 and n_2 are positive integers, and $a_{ij}^{cd} > 0$ is the DAC edge weight of $(i,j) \in \mathscr{E}_c$. Moreover, $\eta_{i,j}^c(t), \widehat{V}_{i,j}^c(t)$ denote the required information from DGU j, and matrices $A_1, B_1, C_1, A_2, B_2, C_2$ and scalar $\gamma > 0$ are DAC parameters invariant among all DGUs. The output of (22b), i.e., $\eta_i(t)$, acts as the feedback signal between input $V_i(t)$ and output $\widehat{V}_i(t)$. The DAC estimator (22) achieves **Robust Average Consensus** (RAC) if $\widehat{V}_i(t)$ tracks the APV with zero steady-state error, i.e.,

$$\widehat{V}_i(\infty) - \langle \boldsymbol{v}(\infty) \rangle = 0, \tag{23}$$

regardless of the initial internal states $\boldsymbol{X}_{i1}(0), \boldsymbol{X}_{i2}(0), \forall i \in \mathscr{V}$. It is worth noting that RAC plays a vital role in supporting the plugging-in/out operations in DCmGs, as these operations will inevitably incur nonzero initial internal states for DAC estimators. Referring to Theorem 2 of [24], we obtain the following result for (22).

Lemma 1: Under Assumption 3 and PCC voltages satisfying $V_i(s) = \frac{c_i^c(s)}{s} + \frac{c_i^r(s)}{s^2}, \forall i \in \mathscr{V}$, [3] where $c_i^c(s)$ and $c_i^r(s)$ are polynomials that may differ among DGUs, all DAC estimators (22) in the DCmG can achieve RAC if

$$h(s) = C_1 (s\mathrm{I}^{n_1} - A_1)^{-1} B_1 = \frac{2as + a^2}{(s+a)^2}, \tag{24a}$$

$$g(s) = C_2 (s\mathrm{I}^{n_2} - A_2)^{-1} B_2 = \frac{s+a}{s^2}, \tag{24b}$$

where $a > 0$ is an arbitrary scalar. Moreover, the minimal realizations for $h(s)$ and $g(s)$ are utilized, i.e.,

$$A_1 = \begin{bmatrix} -2a & -a^2 \\ 1 & 0 \end{bmatrix}, B_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, C_1 = \begin{bmatrix} 2a & a^2 \end{bmatrix}, \tag{25a}$$

$$A_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ C_2 = \begin{bmatrix} 1 & a \end{bmatrix}. \tag{25b}$$

Proof: The proof is provided in Appendix E of the supplementary material. ■

Remark 7: The statement in Lemma 1 also holds when $V_i(s) = \frac{c_i^c(s)}{q_i^c(s)s} + \frac{c_i^r(s)}{q_i^r(s)s^2}, \forall i \in \mathscr{V}$, where $q_i^c(s), q_i^r(s)$ are stable polynomials with all roots lying in the open left half-plane and contribute exponentially vanishing components to $V_i(s)$. According to Theorems 2-3, under ZTS attacks with constant $\bar{\boldsymbol{d}}_{ij}^a, (i,j) \in \bar{\mathscr{E}}_c$, PCC voltages will eventually converge to stable values ($V_i(s) = \frac{c_i^c(s)}{q_i^c(s)s}$) or grow like ramp signals ($V_i(s) = \frac{c_i^r(s)}{q_i^r(s)s^2}$). Hence, under ZTS attacks satisfying Assumption 5, all DAC estimators (22) with parameters set as (25) can achieve RAC. Here we consider the minimal realizations for $h(s)$ and $g(s)$ as they require the minimum number of the internal states in (22).

To ensure the integrity of the DAC related information communicated between DGUs i and j, the UIO-based detectors are deployed. Specifically, DGU i will utilize the received output information from DGU j to estimate the internal states of the DAC estimator, and then compute residuals to detect possible attacks. Nevertheless, in terms of the estimator dynamics (22) with parameters set as (25), the UIO-based detectors are unable to detect any attack due to $\mathrm{rank}(B_1) = \mathrm{rank}(C_1)$ and $\mathrm{rank}(B_2) = \mathrm{rank}(C_2)$ [18]. That is, the number of decoupled unknown inputs is equal to the number of received independent outputs, and thus FDI attacks on the communicated outputs are indistinguishable from those caused by unknown inputs. To enable the attack detection ability of the UIO-based detectors, DGU j will transmit the internal states $\boldsymbol{X}_{j1}(t)$ and $\boldsymbol{X}_{j2}(t)$ to DGU i, such that the number of received independent

[3] $V_i(s)$ denotes the Laplace transform of $V_i(t)$.

outputs can be increased. Moreover, the performance of the DAC estimator in tracking the APV will not be affected as the required information $\widehat{V}_{i,j}^c(t)$ and $\eta_{i,j}^c(t)$ can be calculated from the received internal states $\boldsymbol{X}_{i,j1}^c(t)$ and $\boldsymbol{X}_{i,j2}^c(t)$, respectively.

Lemma 2: The integrity of the communicated data $\boldsymbol{X}_{i,j1}^c(t)$ and $\boldsymbol{X}_{i,j2}^c(t)$ is guaranteed via the following UIO-based detectors, i.e.,

$$\mathrm{UIO}_{i,j}^v \begin{cases} \dot{\boldsymbol{z}}_{i,j}^v(t) = F_j^v \boldsymbol{z}_{i,j}^v(t) + \hat{K}_j^v \boldsymbol{X}_{i,j1}^c(t), \\ \widehat{\boldsymbol{X}}_{i,j1}(t) = \boldsymbol{z}_{i,j}^v(t) + H_j^v \boldsymbol{X}_{i,j1}^c(t), \end{cases} \tag{26a}$$

$$\mathrm{UIO}_{i,j}^\eta \begin{cases} \dot{\boldsymbol{z}}_{i,j}^\eta(t) = F_j^\eta \boldsymbol{z}_{i,j}^\eta(t) + \hat{K}_j^\eta \boldsymbol{X}_{i,j2}^c(t), \\ \widehat{\boldsymbol{X}}_{i,j2}(t) = \boldsymbol{z}_{i,j}^\eta(t) + H_j^\eta \boldsymbol{X}_{i,j2}^c(t), \end{cases} \tag{26b}$$

such that the detection residuals

$$\boldsymbol{r}_{i,j1}^v(t) = \boldsymbol{X}_{i,j1}^c(t) - \widehat{\boldsymbol{X}}_{i,j1}(t) = e^{F_j^v t}\boldsymbol{\epsilon}_{i,j1}^v(0), \tag{27a}$$

$$\boldsymbol{r}_{i,j2}^\eta(t) = \boldsymbol{X}_{i,j2}^c(t) - \widehat{\boldsymbol{X}}_{i,j2}(t) = e^{F_j^\eta t}\boldsymbol{\epsilon}_{i,j2}^\eta(0), \tag{27b}$$

both decay exponentially to zero in the absence of attacks. Here the UIO parameters $F_j^v, \hat{K}_j^v, H_j^v$ and $F_j^\eta, \hat{K}_j^\eta, H_j^\eta$ are set according to (8) and (25) to ensure that F_j^v, F_j^η are both Hurwitz stable. Moreover, $\boldsymbol{\epsilon}_{i,j1}^v(0)$ and $\boldsymbol{\epsilon}_{i,j2}^\eta(0)$ are the initial state estimation errors.

Proof: The proof is provided in Appendix F of the supplementary material. ∎

Similar to (10), the time-varying detection thresholds can be calculated such that

$$|\boldsymbol{r}_{i,j1}^v(t)| \le \bar{\boldsymbol{r}}_{i,j1}^v(t) = \kappa^v e^{-\mu^v t}\bar{\boldsymbol{\epsilon}}_{i,j1}^v(0), \tag{28a}$$

$$|\boldsymbol{r}_{i,j2}^\eta(t)| \le \bar{\boldsymbol{r}}_{i,j2}^\eta(t) = \kappa^\eta e^{-\mu^\eta t}\bar{\boldsymbol{\epsilon}}_{i,j2}^\eta(0) \tag{28b}$$

always hold in the absence of attacks. Once (28a) or (28b) is violated, it is considered that the received $\boldsymbol{X}_{i,j1}^c(t)$ or $\boldsymbol{X}_{i,j2}^c(t)$ from DGU $j \in \mathcal{N}_i^c$ is corrupted.

Nevertheless, the attacker is still able to construct ZTS-like attacks to bypass the UIO-based detectors (26), once he/she has full knowledge of the DAC parameters $A_1, B_1, C_1, A_2, B_2, C_2$, which are completely determined by the scalar a. According to the attack model, the attacker can obtain some system parameters including a from insiders every few hours or days, while the real-time access to a is infeasible. Hence, we attempt to conceal a from the attacker based on the moving target defense (MTD) strategy, whose basic idea is to proactively perturb system parameters to make the attacker's understanding of system model outdated [25].

Assumption 6: The DAC parameters $A_1, B_1, C_1, A_2, B_2, C_2$ can be concealed from the attacker based on the MTD strategy.

Remark 8: To conceal the DAC parameters from the attacker, the perturbation strategy should be designed such that 1) the attacker cannot obtain the explicit perturbation command on a, which is denoted by Δa; 2) the attacker cannot infer Δa from available information immediately. The first objective can be achieved by transmitting Δa through encryption-based secure channels. While the second objective requires to perturb some extra parameters besides a as the identification of transfer functions $h(s)$ and $g(s)$ (i.e., a) is possible after collecting enough inputs and outputs of the two linear dynamical systems involved in (22). We choose to additionally perturb the DAC edge weights $a_{ij}^{cd}, \forall (i,j) \in \mathscr{E}$ to hinder the identification of $h(s)$ and $g(s)$, as the inference of a_{ij}^{cd} is usually time-consuming [26]. Therefore, if the control center can transmit the perturbation commands on a and a_{ij}^{cd} to all DGUs through secure channels every 5/10 minutes, then Assumption 6 could be achieved. Moreover, it is noted that the perturbation will not impact the RAC once $a > 0$ and $a_{ij}^{cd} > 0$ are guaranteed, which is validated in Appendix H of the supplementary material.

B. Attack Detection Phase

In this subsection, we introduce the detection indicator and the corresponding detection threshold, under which the detectability for ZTS attacks is investigated. By comparing V_{ref} with $\widehat{V}_i(t)$, we obtain the estimated average PCC voltage deviation (APVD) as

$$\widehat{V}_i^{err}(t) = V_{ref} - \widehat{V}_i(t). \tag{29}$$

Although the daily operations (e.g., load switches and plugging-in/out of DGUs) in DCmGs never cause steady-state APVD, i.e., $\langle \boldsymbol{v}(\infty) \rangle = V_{ref}$, non-trivial instantaneous APVD will emerge as it takes some time for $\langle \boldsymbol{v}(t) \rangle$ to converge as Fig. 3 shows. Thus, both daily operations and ZTS attacks can result in non-trivial $\widehat{V}_i^{err}(t)$, and it is difficult to distinguish attacks from daily operations based on only historical and current non-trivial $\widehat{V}_i^{err}(t)$. Fortunately, we observe that the non-trivial $|\widehat{V}_i^{err}(t)|$ caused by daily operations shares one common characteristic, i.e., $|\widehat{V}_i^{err}(t)|$ reaches the peak value at almost the time when daily operations occur and then it will quickly decay to zero. Differently, under Assumption 5, the ZTS attack (12) will cause either constant or ramp-growing APVD. Thus, it is natural to derive the following detection indicator based on the sliding time window (STW) technology.

Definition 4 (STW-based Detection Indicator): Given the time window with fixed length T, the detection indicator $\mathscr{A}_i(t)$ is computed as the integral of the time window $(t-T, t)$ sliding over $|\widehat{V}_i^{err}(t)|$, i.e.,

$$\mathscr{A}_i(t) = \begin{cases} 0, t_s + T > t \ge t_s, \\ \displaystyle\int_{t-T}^{t} |\widehat{V}_i^{err}(\tau)| \mathrm{d}\tau, t \ge t_s + T, \end{cases} \tag{30}$$

where $t_s > 0$ is the activation time for the generation of $\mathscr{A}_i(t)$.

Next, we investigate the setting of the detection threshold under which certain daily operations can be tolerated. Let $\mathscr{O}(t) = \{o_1(t), \cdots, o_{|\mathscr{O}|}(t)\}$ be the set of daily operations, where $o_k(t) \in \mathscr{O}(t)$ represents the event of a daily operation occurring at time t. To tolerate any daily operation contained in $\mathscr{O}(t), \forall t \ge t_s$, the constant detection threshold is set as

$$\bar{\mathscr{A}}_i = \max_{\substack{t \ge t_s + T \\ o_k(t_s) \in \mathscr{O}(t_s)}} \int_{t-T}^{t} |\widehat{V}_{i|o_k(t_s)}^{err}(\tau)| \mathrm{d}\tau, t \ge t_s + T, \tag{31}$$

where $\widehat{V}_{i|o_k(t_s)}^{err}(t)$ denotes the estimated APVD under the daily operation $o_k(t_s)$, and could be obtained from the historical real world data or the simulated data. To preserve the detectability

for ZTS attacks, it is suggested to tolerate the most frequent daily operations in DCmGs. Thus, under any daily operation $o_k(t) \in \mathcal{O}(t), \forall t \geq t_s$, we have

$$\mathscr{d}_i(t) \leq \bar{\mathscr{d}}_i, \forall t \geq t_s. \tag{32}$$

If (32) is violated, it is considered that there exist ZTS attacks and the impact mitigation phase will be activated.

Theorem 4: Under Assumptions 2-6 and the DAC estimators (22) satisfying (25), ZTS attacks (12) can be detected by the STW-based detection indicator $\mathscr{d}_i(t)$ if

$$\langle \boldsymbol{\psi}_a(\infty) \rangle > \frac{1}{T} \bar{\mathscr{d}}_i, \tag{33}$$

where $\langle \boldsymbol{\psi}_a(t) \rangle$ denotes the APVD caused by ZTS attacks.

Proof: The proof is given in Appendix C. ■

Remark 9: Here the detection threshold $\bar{\mathscr{d}}_i$ does not increase linearly with the time window length T, as the estimated APVD will eventually converge to zero, i.e., $\widehat{V}^{err}_{i|o_k(t_s)}(\infty) = 0$. Thus, according to (33), the detectability for the ZTS attacks with constant $\bar{\boldsymbol{d}}^a_{ij}$ would be enhanced with a larger T. That is, a larger T can decrease the impact of daily operations on the detectability for ZTS attacks. Meanwhile, we should also be aware of that the STW technology will result in certain amount of initial detection delay and some computation burden for each DGU, and thus T cannot be set arbitrarily large.

C. Impact Mitigation Phase

In this subsection, we introduce the impact mitigation phase that will be activated once (32) is violated. In particular, let $t = T^{alm}_i$ be the time when $\mathscr{d}_i(T^{alm}_i) > \bar{\mathscr{d}}_i$, after which the estimated APVD $\widehat{V}^{err}_i(t)$ is fed into the proportional-integral (PI) based compensator, i.e.,

$$C_i(t) = k_{cp} \widehat{V}^{err}_i(t) + k_{ci} \int_{T^{alm}_i}^{t} \widehat{V}^{err}_i(\tau) \mathrm{d}\tau, \tag{34}$$

where $k_{cp} > 0$ and $k_{ci} > 0$ are PI compensation gains. The compensation signal $C_i(t)$ will be added to the secondary control input, and with some abuse of the notation, the compensated secondary control input is denoted by $\psi_i(t)$, i.e.,

$$\psi_i(t) = \tilde{\psi}_i(t) + \psi^a_i(t) + C_i(t), \tag{35}$$

where $\tilde{\psi}_i(t)$ is the healthy component and $\psi^a_i(t)$ is the malicious component associated with attacks.

Theorem 5: Under Assumptions 2-6 and the DAC estimators (22) satisfying (25), if (32) is violated, then the activated impact mitigation strategy (35) can eventually achieve

$$\langle \boldsymbol{v}(\infty) \rangle = V_{ref} - \frac{1}{k_{ci}} \sum_{(i,j) \in \bar{\mathcal{E}}_c} \frac{k_I a^c_{ij}}{N I^s_{tj}} \boldsymbol{k}^{\mathrm{T}} A^{-1}_{kj} \bar{E}_j \bar{\boldsymbol{d}}^a_{ij}. \tag{36}$$

Proof: The proof is given in Appendix D. ■

Remark 10: For the cooperative ZTS attacks satisfying (20), the impact mitigation strategy (35) can eliminate the constant APVD caused by them. Regarding to the non-cooperative attacks where (20) is not satisfied and ramp-growing APVD is induced, the impact mitigation strategy (35) can stabilize all PCC voltages with constant APVD, which is determined by (36). Moreover, we note that the impact mitigation strategy (35) will not destroy voltage balancing when some daily operations falsely trigger the attack alarm (32) (i.e., the false alarm) once the PI compensation gains are well tuned, which is validated in Appendix I of the supplementary material.

If the detection indicator $\mathscr{d}_i(t)$ is smaller than a predefined threshold $T\delta > 0$ at $t = T^{com}_i$, i.e.,

$$|\mathscr{d}_i(T^{com}_i)| < T\delta, \tag{37}$$

it is considered that voltage balancing has almost been recovered. Then, the impact mitigation strategy (35) is disabled, and the corresponding compensation value $C_i(T^{com}_i)$ is added to the secondary control input as a constant.

Although we only analyze the effectiveness of Algorithm 1 under the ZTS attacks with constant $\bar{\boldsymbol{d}}^a_{ij}$, it is emphasized that Algorithm 1 is also effective when $\bar{\boldsymbol{d}}^a_{ij}(t)$ is time-varying. In simulations, we show the effectiveness of Algorithm 1 under the ZTS attack with sine signal $\bar{\boldsymbol{d}}^a_{ij}(t)$ in Fig. 8.

V. Simulation Studies

In this section, we conduct extensive simulation studies on the DCmG composed of 8 DGUs established in Matlab Simulink/PLECS to validate the theoretical results. The corresponding electrical parameters are provided in Appendix G of the supplementary material. The nominal reference voltage is set as $V_{ref} = 48\text{V}$, and the bounds of noises are $\bar{\boldsymbol{\rho}}_i = \bar{\boldsymbol{\omega}}_i = [0.001, 0.003, 0]^{\mathrm{T}}, \forall i \in \mathscr{V} = \{1, \cdots, 8\}$. The weight parameter involved in (3) is $k_I = 5$. The DAC parameters are set according to (25) with $a = 100$, and the length of STW is $T = 0.65\text{s}$. Moreover, the PI compensation gains in (34) are $k_{cp} = 1, k_{ci} = 20$, and the threshold judging the achievement of voltage balancing is set as $\delta = 0.005\text{V}$.

A. The Setting of $\bar{\mathscr{d}}_i$

In this subsection, we investigate the setting of detection thresholds $\bar{\mathscr{d}}_i, \forall i \in \mathscr{V}$ that can tolerate any daily operation contained in the set $\mathcal{O}(t) = \{o_1(t), o_2(t), o_3(t)\}$, whose elements are elaborated in TABLE I. Before implementing $\mathcal{O}(t)$, a series of initialization operations are conducted as indicated by Fig. 2, which are detailed as follows: at $t = 0\text{s}$, all primary controllers are activated; at $t = 2\text{s}$, all DGUs except DGU 7 are connected through power lines; at $t = 4\text{s}$, the communication network is established, and UIO-based detectors (7), (26), DAC estimators (22), and the generation of $\mathscr{d}_i(t)$ are activated. Then, the daily operations are introduced: at $t = 8\text{s}$, DGU 8 is plugged out from the DCmG; at $t = 12\text{s}$, all load currents are decreased by 30% of their rated values; at $t = 16\text{s}$, DGUs 7 is plugged into the DCmG.

TABLE I
Elaboration of Daily Operation Set $\mathcal{O}(t)$

$o_1(t)$	plugging out of DGU 8
$o_2(t)$	decrease of load currents by 30% of their rated values
$o_3(t)$	plugging in of DGU 7

As shown in Fig. 3, each daily operation will cause non-trivial $|\widehat{V}^{err}_i(t)|$, and some fluctuation emerges on the detection indicator $\mathscr{d}_i(t)$ accordingly. Moreover, it is observed that

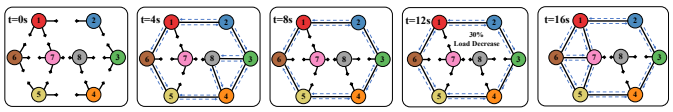


Fig. 2. This figure shows the evolution of the DCmG corresponding to the daily operation set $\mathcal{O}(t)$. Here the solid black lines are power lines and the dotted black lines signify communication links.

any daily operation in $\mathcal{O}(t)$ can be tolerated by the detection thresholds $\bar{d}_i = 0.0325, \forall i \in \mathcal{V}$. According to Theorem 4, the ZTS attacks causing steady-state APVD more than $\frac{\bar{d}_i}{T} = 0.05$V can be detected.

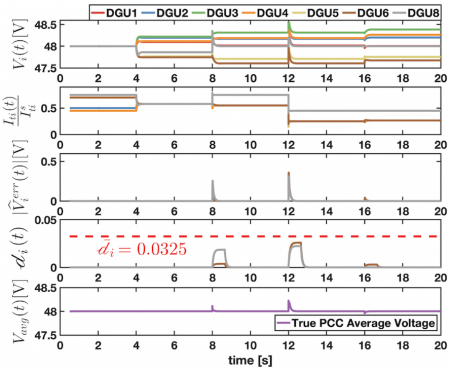
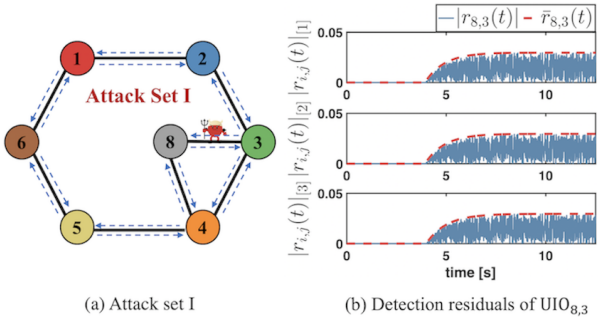


Fig. 3. This figure shows PCC voltages $V_i(t)$, output currents in per-unit $\frac{I_{ti}(t)}{I_{ti}^s}$, estimated APVDs $\widehat{V}_i^{err}(t)$, detection indicators $d_i(t), \forall i \in \mathcal{V}$ and the true average PCC voltage $V_{avg}(t)$.

B. ZTS Attacks with Constant $\bar{d}_{ij}^a$

In this subsection, we validate the effectiveness of Algorithm 1 against the ZTS attacks with constant $\bar{d}_{ij}^a$. In particular, we consider two cases where single ZTS attack and cooperative ZTS attacks are launched.

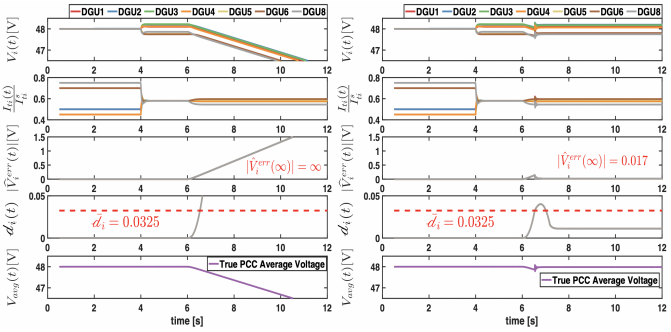
1) Attack Set I*:* Attack set I is composed of one ZTS attack targeting at communication link $(8, 3)$, and the attack vector is generated by (12) with $\bar{d}_{83}^{a1} = [2, 0]^{\mathrm{T}}$. Attack set I is activated at $T_{a1} = 6$s. As shown in (b) of Fig. 4, the detection residuals under attack set I are still bounded by the detection thresholds, indicating that attack set I is unforeseeable for $\mathrm{UIO}_{8,3}$.



(a) Attack set I (b) Detection residuals of $\mathrm{UIO}_{8,3}$

Fig. 4. This figure depicts the location of attack set I and the detection residuals of $\mathrm{UIO}_{8,3}$ under attack set I.

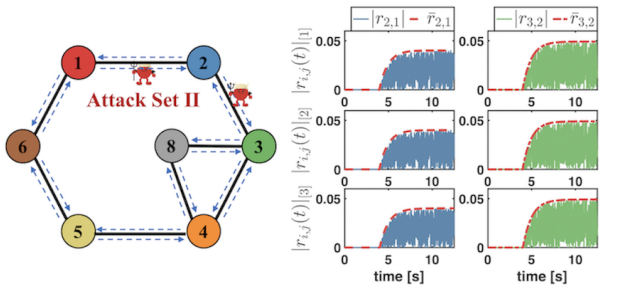
According to Fig. 5, attack set I incurs ramp-growing APVD and current sharing is damaged. Obviously, the activated countermeasure can make the PCC voltages finally converge and the steady-state APVD is 0.017V, which can significantly mitigate the attack impact and timely avoid the occurrence of a blackout incident in the DCmG.



(a) Without compensation (b) With compensation

Fig. 5. This figure shows PCC voltages, output currents and the countermeasure related variables under attack set I without compensation ($k_{cp} = k_{ci} = 0$) and with compensation ($k_{cp} = 1, k_{ci} = 20$).

2) Attack Set II*:* Attack set II is composed of two cooperative ZTS attacks targeting at communication links $(2, 1)$ and $(3, 2)$, and the corresponding attack vectors are generated by (12) with parameters $\bar{d}_{21}^{a2} = [2, 0]^{\mathrm{T}}$ and $\bar{d}_{32}^{a2} = [-2.8, 0]^{\mathrm{T}}$, respectively. Attack set II is activated at $T_{a2} = 6$s. Similarly, as shown in (b) of Fig. 6, attack set II can bypass the detection of $\mathrm{UIO}_{2,1}$ and $\mathrm{UIO}_{3,2}$, as the corresponding detection residuals are almost not impacted.



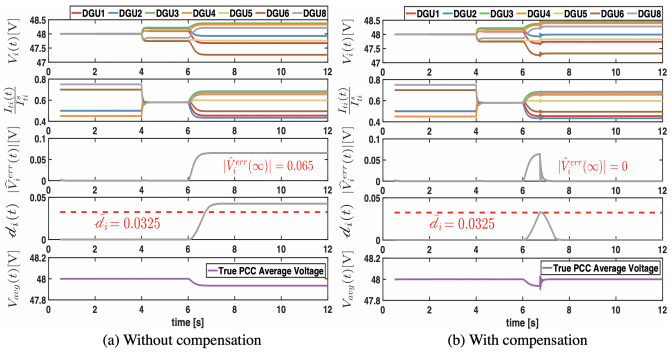
(a) Attack set II (b) Detection residuals of $\mathrm{UIO}_{2,1}$ and $\mathrm{UIO}_{3,2}$

Fig. 6. This figure depicts the location of attack set II and the detection residuals of $\mathrm{UIO}_{2,1}$ and $\mathrm{UIO}_{3,2}$ under attack set II.

According to Fig. 7, attack set II causes constant APVD and destroys current sharing in DCmGs, which validates the correctness of Theorem 3. The activated countermeasure can effectively eliminate the malicious APVD and pull the PCC voltages around the nominal reference point, which validates the statement in Remark 10.

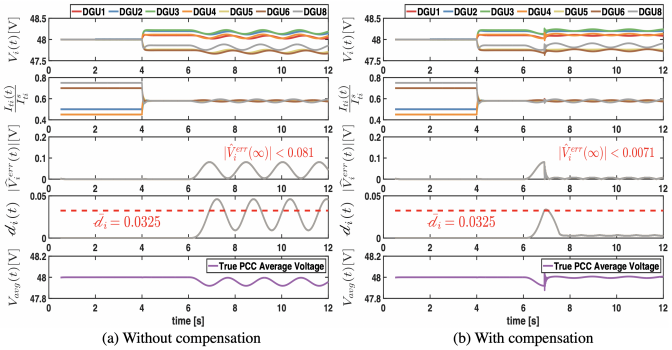
C. ZTS Attack with Time-varying $\bar{d}_{ij}^a(t)$

In this subsection, the effectiveness of Algorithm 1 against the ZTS attacks with time-varying $\bar{d}_{ij}^a(t)$ is shown. Attack set III is composed of one ZTS attack targeting at communication link $(8, 3)$, and the attack vector is generated by (12) with $\bar{d}_{83}^{a3}(t) = [\sin(4t), 0]^{\mathrm{T}}$. Attack set III is activated at $T_{a3} = 6$s. According to Fig. 8, it is validated that the countermeasure



(a) Without compensation (b) With compensation

Fig. 7. This figure shows PCC voltages, output currents and countermeasure related variables under attack set II without compensation ($k_{cp} = k_{ci} = 0$) and with compensation ($k_{cp} = 1, k_{ci} = 20$).



(a) Without compensation (b) With compensation

Fig. 8. This figure shows PCC voltages, output currents and countermeasure related variables under under attack set III without compensation ($k_{cp} = k_{ci} = 0$) and with compensation ($k_{cp} = 1, k_{ci} = 20$).

can substantially decrease the APVD caused by attack set III, such that the APVD after compensation is neglectable.

For clarity, we present TABLE II to sum up the APVDs without compensation and with compensation under the three attack sets aforementioned.

TABLE II
APVDs UNDER THE THREE ATTACK SETS

	Without compensation	With compensation
Attack set I	$\lvert\widehat{V}_i^{err}(\infty)\rvert = \infty$V	$\lvert\widehat{V}_i^{err}(\infty)\rvert = 0.017$V
Attack set II	$\lvert\widehat{V}_i^{err}(\infty)\rvert = 0.065$V	$\lvert\widehat{V}_i^{err}(\infty)\rvert = 0$V
Attack set III	$\lvert\widehat{V}_i^{err}(\infty)\rvert \leq 0.081$V	$\lvert\widehat{V}_i^{err}(\infty)\rvert \leq 0.0071$V

VI. Conclusion

In this paper, we revealed that the vulnerability of the UIO-based detector originates from the lack of knowledge of true inputs, and ZTS attacks can be constructed by secretly faking the unknown inputs. Moreover, it was proved that single ZTS attack can destabilize the DCmG, and cooperative ZTS attacks can cause accurate and specific impact. Based on the estimated APV, we designed a distributed countermeasure against ZTS attacks, which can decrease the APVD or even recover voltage balancing in DCmGs. In [27], we have proposed a converter-based moving target defense strategy to eliminate the threat of ZTS attacks, where the primary control gains are proactively perturbed to invalidate the attacker's understanding of system parameters. In the future work, we will rigorously investigate the PI compensation gains' stability region under which the extreme daily operations will not destabilize the PCC voltages.

Appendix

A. Proof of Theorem 2

After simplifying the primary control loops as unit gains [6], we have

$$\boldsymbol{v}(t) = \boldsymbol{v}_r + \boldsymbol{\psi}(t), \tag{38}$$

where $\boldsymbol{v}_r = V_{ref}\mathbb{1}^N$ is the constant vector containing nominal reference PCC voltages. Moreover, integrating (3) with (4), the dynamics of the secondary control input under attacks are

$$\dot{\boldsymbol{\psi}}(t) = -\tilde{L}D\boldsymbol{i}_t(t) + \frac{k_I a_{ij}^c}{I_{tj}^s}\boldsymbol{k}^{\mathrm{T}}\boldsymbol{\phi}_{i,j}(t)\boldsymbol{l}_i, \tag{39}$$

where $\tilde{L} = k_I L, D = \mathrm{diag}\{\frac{1}{I_{t1}^s}, \cdots, \frac{1}{I_{tN}^s}\}$, and $\boldsymbol{l}_i \in \mathbb{R}^N$ is obtained from $\mathbb{0}^N$ with its i-th element replaced by 1. Here $\boldsymbol{i}_t(t) = [I_{t1}(t), \cdots, I_{tN}(t)]^{\mathrm{T}}$ is the output current vector under attacks and according to the Kirchhoff current law, we obtain

$$\boldsymbol{i}_t(t) = M\boldsymbol{v}(t) + \boldsymbol{i}_l, \tag{40}$$

where $\boldsymbol{i}_l = [I_{L1}, \cdots, I_{LN}]^{\mathrm{T}}$ is the constant vector containing load currents. The overall dynamics of DCmGs can be obtained after integrating equations (38)-(40), i.e.,

$$\dot{\boldsymbol{\psi}}(t) = -Q\boldsymbol{\psi}(t) - \tilde{L}D\boldsymbol{i}_l - Q\boldsymbol{v}_r + \frac{k_I a_{ij}^c}{I_{tj}^s}\boldsymbol{k}^{\mathrm{T}}\boldsymbol{\phi}_{i,j}(t)\boldsymbol{l}_i, \tag{41}$$

where $Q = \tilde{L}DM$ integrates the Laplacian matrices of graphs $\mathcal{G}_c$ and $\mathcal{G}_{el}$. According to Proposition 3 in [6], Q satisfies a) $\ker(Q)$=$\mathbb{H}^1_\perp$, range(Q)= $\mathbb{H}^1$ and b) Q is diagonalizable and has non-negative eigenvalues, and its algebraic multiplicity of zero eigenvalue is one.

Hence, pairs containing eigenvalues and eigenvectors of Q are denoted by $\boldsymbol{p}_i = (\lambda_i, \boldsymbol{q}_i), \forall i \in \mathscr{V}$, where $\lambda_1 = 0$, $0 < \lambda_2 \leq \cdots \leq \lambda_N$, $\boldsymbol{q}_i \in \mathbb{R}^N$, $\boldsymbol{q}_1 \in \mathbb{H}^1_\perp$, and the set $\{\boldsymbol{q}_2, \cdots, \boldsymbol{q}_N\}$ constitutes a basis of $\mathbb{H}^1$. Given the linear differential equation (41), the healthy component of the secondary control input vector can be decomposed and calculated as

$$\tilde{\boldsymbol{\psi}}(t) = e^{-Qt}\tilde{\boldsymbol{\psi}}(0) + \sum_{i=2}^{N}\frac{\alpha_i}{\lambda_i}(1 - e^{-\lambda_i t})\boldsymbol{q}_i, \tag{42}$$

where $\alpha_i \in \mathbb{R}, \forall i \in \mathscr{V}$ are chosen such that $\sum_{i=2}^N \alpha_i \boldsymbol{q}_i = -\tilde{L}D\boldsymbol{i}_l - Q\boldsymbol{v}_r \in \mathbb{H}^1$.

Under Assumption 5, the ZTS attack vector can be expressed as

$$\boldsymbol{\phi}_{i,j}(t) = e^{A_{kj}(t-T_a)}\check{\boldsymbol{\phi}}_{i,j}(T_a) - A_{kj}^{-1}\bar{E}_j\bar{\boldsymbol{d}}_{ij}^a, \tag{43}$$

where $\check{\boldsymbol{\phi}}_{i,j}(T_a) = A_{kj}^{-1}\bar{E}_j\bar{\boldsymbol{d}}_{ij}^a$. The malicious component of the secondary control input vector is decomposed as $\boldsymbol{\psi}_a(t) = \boldsymbol{\psi}_{a1}(t) + \boldsymbol{\psi}_{a2}(t)$, where $\boldsymbol{\psi}_{a1}(t)$ and $\boldsymbol{\psi}_{a2}(t)$ represent the components associated with $e^{A_{kj}(t-T_a)}\check{\boldsymbol{\phi}}_{i,j}(T_a)$ and $-A_{kj}^{-1}\bar{E}_j\bar{\boldsymbol{d}}_{ij}^a$, respectively. As $A_{kj}^{-1}\bar{E}_j\bar{\boldsymbol{d}}_{ij}^a$ is a constant vector, $\boldsymbol{\psi}_{a2}(t)$ can be directly calculated as

$$\boldsymbol{\psi}_{a2}(t) = \alpha_{1l}(t - T_a)\boldsymbol{q}_1 + \sum_{i=2}^{N}\frac{\alpha_{il}}{\lambda_i}(1 - e^{-\lambda_i(t-T_a)})\boldsymbol{q}_i, \tag{44}$$

where $\alpha_{il} \in \mathbb{R}, \forall i \in \mathscr{V}$ are chosen such that $\sum_{i=1}^N \alpha_{il}\boldsymbol{q}_i = -\frac{k_I a_{ij}^c}{I_{tj}^s}\boldsymbol{k}^{\mathrm{T}}A_{kj}^{-1}\bar{E}_j\bar{\boldsymbol{d}}_{ij}^a\boldsymbol{l}_i$. For brevity, we only show the

case where $\boldsymbol{A_{kj}}$ is diagonalizable[4], with which we have $e^{A_{kj}(t-T_a)}\check{\boldsymbol{\phi}}_{i,j}(T_a) = \sum_{m=1}^{3}\eta_m e^{\beta_m(t-T_a)}\boldsymbol{a}_m$. Here, pairs $(\beta_m, \boldsymbol{a}_m), \forall m \in \underline{\mathscr{V}} = \{1,2,3\}$ contain the eigenvalues and eigenvectors of A_{kj}, and $\eta_m, \forall m \in \underline{\mathscr{V}}$ are scalars such that $\sum_{m=1}^{3}\eta_m \boldsymbol{a_m} = \check{\boldsymbol{\phi}}_{i,j}(T_a)$. Then, $\boldsymbol{\psi}_{a1}(t)$ is calculated as

$$\boldsymbol{\psi}_{a1}(t) = \sum_{m=1}^{3}\frac{k_I a_{ij}^c}{I_{tj}^s}\eta_m a_{m2} e^{-\beta_m T_a} e^{-Qt}\int_{T_a}^{t} e^{\beta_m \tau} e^{Q\tau}\boldsymbol{l}_i d\tau, \quad (45)$$

where $a_{m2} = \boldsymbol{k}^{\mathrm{T}}\boldsymbol{a}_m \in \mathbb{R}$ denotes the second entry of vector $\boldsymbol{a}_m$. As the set $\{\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_N\}$ constitutes a basis of $\mathbb{R}^N$, $\boldsymbol{l}_i$ can be expressed as $\boldsymbol{l}_i = \sum_{r=1}^{N}\delta_r \boldsymbol{q}_r$, where $\delta_1 \boldsymbol{q}_1 = \langle \boldsymbol{l}_i \rangle \mathbb{1}^N = \frac{1}{N}\mathbb{1}^N$. Thus, we obtain $e^{Qt}\boldsymbol{l}_i = \sum_{r=1}^{N}\delta_r e^{\lambda_r t}\boldsymbol{q}_r$. Similarly, we only show the case where $\beta_m + \lambda_r \neq 0, \forall m \in \underline{\mathscr{V}}, r \in \mathscr{V}$,[5] then the integral component in (45) is computed as

$$\int_{T_a}^{t} e^{\beta_m \tau} e^{Q\tau}\boldsymbol{l}_i d\tau = \sum_{r=1}^{N}\frac{\delta_r}{\beta_m + \lambda_r}(e^{(\beta_m+\lambda_r)t} - e^{(\beta_m+\lambda_r)T_a})\boldsymbol{q}_r. \quad (46)$$

Substituting (46) into (45), we have

$$\boldsymbol{\psi}_{a1}(t) = \sum_{m=1}^{3}\sum_{r=1}^{N}\frac{k_I a_{ij}^c \eta_m a_{m2}\delta_r}{I_{tj}^s(\beta_m+\lambda_r)}(e^{\beta_m(t-T_a)} - e^{-\lambda_r(t-T_a)})\boldsymbol{q}_r. \quad (47)$$

With $\alpha_{1l}\boldsymbol{q}_1 = -\frac{k_I a_{ij}^c}{NI_{tj}^s}\boldsymbol{k}^{\mathrm{T}}A_{kj}^{-1}\bar{E}_j\bar{\boldsymbol{d}}_{ij}^a\mathbb{1}^N$, the average of elements in $\boldsymbol{\psi}_{a2}(\infty)$ is calculated as

$$\langle\boldsymbol{\psi}_{a2}(\infty)\rangle = \langle\alpha_{1l}\boldsymbol{q}_1\rangle(t-T_a) = -\frac{k_I a_{ij}^c}{NI_{tj}^s}\boldsymbol{k}^{\mathrm{T}}A_{kj}^{-1}\bar{E}_j\bar{\boldsymbol{d}}_{ij}^a(t-T_a). \quad (48)$$

As A_{kj} is Hurwitz stable, i.e., $\mathrm{Re}(\boldsymbol{\beta}_m) < 0, \forall m \in \underline{\mathscr{V}}$, and $\delta_1\boldsymbol{q}_1 = \frac{1}{N}\mathbb{1}^N$, it follows from (47) that

$$\langle\boldsymbol{\psi}_{a1}(\infty)\rangle = -\sum_{m=1}^{3}\frac{k_I a_{ij}^c \eta_m a_{m2}}{NI_{tj}^s\beta_m}. \quad (49)$$

Meanwhile, it is noted that pairs $(\frac{1}{\beta_m}, \boldsymbol{a}_m), \forall m \in \underline{\mathscr{V}}$ includes eigenvalues and eigenvectors of $\boldsymbol{A}_{kj}^{-1}$. Thus, with $e^{A_{kj}(t-T_a)}\check{\boldsymbol{\phi}}_{i,j}(T_a) = \sum_{m=1}^{3}\eta_m e^{\beta_m(t-T_a)}\boldsymbol{a}_m$, we obtain

$$e^{A_{kj}^{-1}}(t-T_a)\check{\boldsymbol{\phi}}_{i,j}(T_a) = \sum_{m=1}^{3}\eta_m e^{\frac{1}{\beta_m}(t-T_a)}\boldsymbol{a}_m. \quad (50)$$

Differentiating both sides of equation (50) and letting $t = T_a$, we have $A_{kj}^{-1}\check{\boldsymbol{\phi}}_{i,j}(T_a) = \sum_{m=1}^{3}\frac{\eta_m}{\beta_m}\boldsymbol{a}_m$, under which (49) is transformed to

$$\langle\boldsymbol{\psi}_{a1}(\infty)\rangle = -\frac{k_I a_{ij}^c}{NI_{tj}^s}\boldsymbol{k}^{\mathrm{T}}A_{kj}^{-1}\check{\boldsymbol{\phi}}_{i,j}(T_a). \quad (51)$$

Integrating (48) with (51), the total attack impact is

$$\langle\boldsymbol{\psi}_a(\infty)\rangle = -\frac{k_I a_{ij}^c}{NI_{tj}^s}\boldsymbol{k}^{\mathrm{T}}A_{kj}^{-1}(\bar{E}_j\bar{\boldsymbol{d}}_{ij}^a(t-T_a) + \check{\boldsymbol{\phi}}_{i,j}(T_a)).$$

Obviously, if $\boldsymbol{k}^{\mathrm{T}}A_{kj}^{-1}\bar{E}_j\bar{\boldsymbol{d}}_{ij}^a$ is not equal to zero, then $\langle\boldsymbol{\psi}_a(t)\rangle$ will diverge with t. Thus, $\langle\boldsymbol{\psi}(t)\rangle$ will also diverge due to $\langle\boldsymbol{\psi}(t)\rangle = \langle\tilde{\boldsymbol{\psi}}(t)\rangle + \langle\boldsymbol{\psi}_a(t)\rangle$, where $\langle\tilde{\boldsymbol{\psi}}(t)\rangle$ decays exponentially to zero as indicated by (42). Thus, the APV $\langle\boldsymbol{v}(t)\rangle$ will grow like the ramp signal, i.e., *voltage balancing is not achieved*. Moreover, with $\boldsymbol{k}^{\mathrm{T}}A_{kj}^{-1}\bar{E}_j\bar{\boldsymbol{d}}_{ij}^a \neq 0$, we have $\langle\frac{k_I a_{ij}^c}{I_{tj}^s}\boldsymbol{k}^{\mathrm{T}}\boldsymbol{\phi}_{i,j}(\infty)\boldsymbol{l}_i\rangle = -\frac{k_I a_{ij}^c}{NI_{tj}^s}\boldsymbol{k}^{\mathrm{T}}A_{kj}^{-1}\bar{E}_j\bar{\boldsymbol{d}}_{ij}^a \neq 0$, indicating that the equilibrium of equation (39) cannot be achieved. Accordingly, given (3), *current sharing cannot be achieved either*. The proof is completed.

[4]When A_{kj} is not diagonalizable, the stability and convergence properties can be analyzed in a similar way.

[5]When $\beta_m = -\lambda_r$, the integral part in (46) can be calculated similarly.

B. Proof of Theorem 3

Given the linear differential equation (41), the impact caused by compromising multi communication links are the sum of the impact that would have been caused by compromising each communication link individually. Then, when cooperative ZTS attack vectors (43) satisfying (20) are injected multi communication links $\tilde{\mathscr{E}}_c$, simultaneously, we obtain

$$\langle\boldsymbol{\psi}_a(\infty)\rangle = -\sum_{(i,j)\in\tilde{\mathscr{E}}_c}\frac{k_I a_{ij}^c}{NI_{tj}^s}\boldsymbol{k}A_{kj}^{-1}(\bar{E}_j\bar{\boldsymbol{d}}_{ij}^a(t-T_a) + \check{\boldsymbol{\phi}}_{i,j}(T_a)). \quad (52)$$

Substituting (20) into (52), we have

$$\langle\boldsymbol{\psi}_a(\infty)\rangle = -\sum_{(i,j)\in\tilde{\mathscr{E}}_c}\frac{k_I a_{ij}^c}{NI_{tj}^s}\boldsymbol{k}A_{kj}^{-1}\check{\boldsymbol{\phi}}_{i,j}(T_a). \quad (53)$$

Similar as the proof of Theorem 2, *voltage balancing will not be achieved if* $\sum_{(i,j)\in\tilde{\mathscr{E}}_c}\frac{k_I a_{ij}^c}{I_{tj}^s}\boldsymbol{k}^{\mathrm{T}}A_{kj}^{-1}\check{\boldsymbol{\phi}}_{i,j}(T_a) \neq 0$, *i.e.,* $\langle\boldsymbol{\psi}_a(\infty)\rangle \neq 0$. Moreover, given (20), we obtain $\langle\sum_{(i,j)\in\tilde{\mathscr{E}}_c}\frac{k_I a_{ij}^c}{I_{tj}^s}\boldsymbol{k}^{\mathrm{T}}A_{kj}^{-1}\bar{E}_j\bar{\boldsymbol{d}}_{ij}^a\boldsymbol{l_i}\rangle = 0$, indicating that the equilibrium of $\boldsymbol{\psi}(t)$ will be achieved as $\langle\tilde{L}D\boldsymbol{i}_t(\infty)\rangle = 0$. To achieve current sharing, i.e., $\tilde{L}D\boldsymbol{i}_t(\infty) = \mathbb{0}^N$, it is necessary to make $\sum_{(i,j)\in\tilde{\mathscr{E}}_c}\frac{k_I a_{ij}^c}{I_{tj}^s}\boldsymbol{k}^{\mathrm{T}}A_{kj}^{-1}\bar{E}_j\bar{\boldsymbol{d}}_{ij}^a\boldsymbol{l}_i = \mathbb{0}^N$. On the contrary, if $\sum_{(i,j)\in\tilde{\mathscr{E}}_c}\frac{k_I a_{ij}^c}{I_{tj}^s}\boldsymbol{k}^{\mathrm{T}}A_{kj}^{-1}\bar{E}_j\bar{\boldsymbol{d}}_{ij}^a\boldsymbol{l}_i \neq \mathbb{0}^N$, then *current sharing cannot be achieved*. The proof is completed.

C. Proof of Theorem 4

Since voltage balancing can always be achieved in the absence of attacks, i.e., $\langle\boldsymbol{\psi}(t)\rangle = 0$, we have

$$\langle\boldsymbol{v}(\infty)\rangle = V_{ref} + \langle\boldsymbol{\psi}_a(\infty)\rangle. \quad (54)$$

Moreover, under Assumption 5, the DAC estimators (22) can always achieve RAC. Accordingly, substituting (54) and (23) into (29), we obtain

$$\widehat{V}_i(\infty) = \langle\boldsymbol{\psi}_a(\infty)\rangle. \quad (55)$$

According to Theorems 2-3, under the ZTS attacks with constant $\bar{\boldsymbol{d}}_{ij}^a, \forall(i,j) \in \tilde{\mathscr{E}}_c$, PCC voltages will either converge to stable values or grow like ramp signals. It is obvious that, if $\langle\boldsymbol{\psi}_a(\infty)\rangle$ is ramp-growing, then the detection indicator $\mathscr{d}_i(t)$ will also keep growing due to (55), indicating that (32) will be eventually violated. Otherwise, if $\langle\boldsymbol{\psi}_a(\infty)\rangle$ is a constant, then with (33), we have

$$\mathscr{d}_i(\infty) = T\langle\boldsymbol{\psi}_a(\infty)\rangle > \bar{\mathscr{d}}_i,$$

which means that (32) is also violated. The state follows.

D. Proof of Theorem 5

First, we consider the cooperative ZTS attacks satisfying (20), and let $\langle\boldsymbol{\psi}_a(\infty)\rangle = -\sum_{(i,j)\in\tilde{\boldsymbol{\varepsilon}}_a}\frac{k_I a_{ij}^c}{NI_{tj}^s}\boldsymbol{k}^{\mathrm{T}}A_{kj}^{-2}\bar{E}_j\bar{\boldsymbol{d}}_{ij}^a = \bar{Q}_c^b$. After activating the impact counteraction strategy (35), the APV is obtained as

$$\langle\boldsymbol{v}(\infty)\rangle = V_{ref} + \bar{Q}_c^b + \langle\boldsymbol{c}(\infty)\rangle, \quad (56)$$

where $\boldsymbol{c}(t)$ collects the compensation values $C_i(t), \forall i \in \mathscr{V}$. Since the DAC estimators (22) can achieve RAC, we have

$$\widehat{V}_i(\infty) = \langle \boldsymbol{v}(\infty) \rangle. \tag{57}$$

Substituting equations (56) and (57) into (29), we obtain

$$\widehat{V}_i^{err}(\infty) = -\bar{Q}_c^b - \langle \boldsymbol{c}(\infty) \rangle. \tag{58}$$

Integrating the PI-based compensator (34) with (58), it is obvious that the equilibrium of (34) will be attained with $\widehat{V}_i^{err}(\infty) = 0$, i.e., voltage balancing is achieved.

Then, we consider the non-cooperative ZTS attacks where (20) is not satisfied, and let $\langle \boldsymbol{\psi}_a(\infty) \rangle = \sum_{(i,j)\in\tilde{\mathscr{E}}_c} -\frac{k_I a_{ij}^c}{N I_{tj}^s} \boldsymbol{k}^{\mathrm{T}} A_{kj}^{-1} (\bar{E}_j \bar{\boldsymbol{d}}_{ij}^a (t - T_a) + A_{kj}^{-1} \bar{E}_j \bar{\boldsymbol{d}}_{ij}^a) = \bar{Q}_n^k t + \bar{Q}_n^b$. Similar to (58), we have

$$\widehat{V}_i^{err}(\infty) = -\bar{Q}_n^k t - \bar{Q}_n^b - \langle \boldsymbol{c}(\infty) \rangle.$$

To track the ramp-growing signal $\bar{Q}_n^k t + \bar{Q}_n^b$ with the PI-based compensator (34), there should be a nonzero tracking error such that $k_{ci} \widehat{V}_i^{err}(\infty) = -\bar{Q}_n^k$. Thus, the APV after compensation can be written as (36), and the proof is completed.

REFERENCES

[1] M. Liu, C. Zhao, R. Deng, P. Cheng, W. Wang, and J. Chen, "Nonzero-dynamics stealthy attack and its impacts analysis in DC microgrids," in *IEEE ACC*. IEEE, 2019, pp. 3922–3927.

[2] N. Hatziargyriou, *Microgrids: architectures and control*. John Wiley & Sons, 2014.

[3] H. Lotfi and A. Khodaei, "AC versus DC microgrid planning," *IEEE Transactions on Smart Grid*, vol. 8, no. 1, pp. 296–304, 2017.

[4] L. Meng, Q. Shafiee, G. F. Trecate, H. Karimi, D. Fulwani, X. Lu, and J. M. Guerrero, "Review on control of DC microgrids and multiple microgrid clusters," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 5, no. 3, pp. 928–948, 2017.

[5] F. Dörfler, J. W. Simpson-Porco, and F. Bullo, "Breaking the hierarchy: distributed control and economic optimality in microgrids," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 3, pp. 241–253, 2015.

[6] M. Tucci, L. Meng, J. M. Guerrero, and G. Ferrari-Trecate, "Stable current sharing and voltage balancing in DC microgrids: a consensus-based secondary control layer," *Automatica*, vol. 95, pp. 1–13, 2018.

[7] H. Zhang, W. Meng, J. Qi, X. Wang, and W. X. Zheng, "Distributed load sharing under false data injection attack in an inverter-based microgrid," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 2, pp. 1543–1551, 2018.

[8] S. Sahoo, S. Mishra, J. C.-H. Peng, and T. Dragičević, "A stealth cyber-attack detection strategy for DC microgrids," *IEEE Transactions on Power Electronics*, vol. 34, no. 8, pp. 8162–8174, 2018.

[9] O. A. Beg, L. V. Nguyen, T. T. Johnson, and A. Davoudi, "Signal temporal logic-based attack detection in DC microgrids," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3585–3595, 2018.

[10] S. Zhao, Q. Yang, P. Cheng, R. Deng, and J. Xia, "Adaptive resilient control for variable-speed wind turbines against false data injection attacks," *IEEE Transactions on Sustainable Energy*, vol. 13, no. 2, pp. 971–985, 2022.

[11] T. M. Chen, "Stuxnet, the real start of cyber warfare?[editor's note]," *IEEE Network*, vol. 24, no. 6, pp. 2–3, 2010.

[12] F. Pasqualetti, F. Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2715–2729, 2013.

[13] S. X. Ding, *Model-based fault diagnosis techniques: design schemes, algorithms, and tools*. Springer Science & Business Media, 2008.

[14] A. Teixeira, H. Sandberg, and K. H. Johansson, "Networked control systems under cyber attacks with applications to power networks," in *IEEE ACC*. IEEE, pp. 3690–3696, 2010.

[15] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "Distributed fault detection and isolation resilient to network model uncertainties," *IEEE Transactions on Cybernetics*, vol. 44, no. 11, pp. 2024–2037, 2014.

[16] M. Tucci, S. Riverso, J. C. Vasquez, J. M. Guerrero, and G. Ferrari-Trecate, "A decentralized scalable approach to voltage control of DC islanded microgrids," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 6, pp. 1965–1979, 2016.

[17] A. Barboni, H. Rezaee, F. Boem, and T. Parisini, "Detection of covert cyber-attacks in interconnected systems: a distributed model-based approach," *IEEE Transactions on Automatic Control*, vol. 65, no. 9, pp. 3728–3741, 2020.

[18] A. J. Gallo, M. S. Turan, F. Boem, T. Parisini, and G. Ferrari-Trecate, "A distributed cyber-attack detection scheme with application to DC microgrids," *IEEE Transactions on Automatic Control*, vol. 65, no. 9, pp. 3800–3815, 2020.

[19] F. Bullo, *Lectures on Network Systems*, 1st ed., 2019.

[20] J. Hunker and C. Probst, "Insiders and insider threats: an overview of definitions and mitigation techniques," *J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl.*, vol. 2, pp. 4–27, 2011.

[21] A. Bidram and A. Davoudi, "Hierarchical structure of microgrids control system," *IEEE Transactions on Smart Grid*, vol. 3, no. 4, pp. 1963–1976, 2012.

[22] M. Crosbie and E. H. Spafford, "Defending a computer system using autonomous agents," *Technical Reporter 95-022, COAST Laboratory-Purdue University*, 1994.

[23] J. Chen, R. J. Patton, and H.-Y. Zhang, "Design of unknown input observers and robust fault detection filters," *International Journal of Control*, vol. 63, no. 1, pp. 85–105, 1996.

[24] H. Bai, R. A. Freeman, and K. M. Lynch, "Robust dynamic average consensus of time-varying inputs," in *IEEE CDC*. IEEE, 2010, pp. 3104–3109.

[25] Z. Zhang, R. Deng, D. K. Yau, P. Cheng, and J. Chen, "Analysis of moving target defense against false data injection attacks on power grid," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 2320–2335, 2019.

[26] S. Segarra, M. T. Schaub, and A. Jadbabaie, "Network inference from consensus dynamics," in *IEEE CDC*. IEEE, 2017, pp. 3212–3217.

[27] M. Liu, C. Zhao, Z. Zhang, R. Deng, P. Cheng, and J. Chen, "Converter-based moving target defense against deception attacks in dc microgrids," *IEEE Transactions on Smart Grid*, pp. 1–1, 2021.

Mengxiang Liu (Member, IEEE) received the B.Sc. degree in automation from Tongji University, Shanghai, in 2017. He is currently pursuing the Ph.D. degree with the College of Control Science and Engineering, Zhejiang University, Hangzhou, China. His research interests include cybersecurity, microgrid, and smart grid.

Chengcheng Zhao (Member, IEEE) received the B.Sc. degree in measurement and control technology and instrument from Hunan University, Changsha, China, in 2013 and the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2018, respectively. She is currently an Associate Researcher in the College of Control Science and Engineering, Zhejiang university. Her research interests include consensus and distributed optimization, distributed energy management and synchronization in smart grids, and security and privacy in networked systems.

Ruilong Deng (Senior Member, IEEE) received the B.Sc. and Ph.D. degrees in control science and engineering from Zhejiang University, Hangzhou, China, in 2009 and 2014, respectively. He is currently a Professor with the College of Control Science and Engineering, Zhejiang University, where he is also with the School of Cyber Science and Technology. His research interests include cybersecurity, smart grid, and wireless networking.

Peng Cheng (Member, IEEE) received the B.Sc. and Ph.D. degrees in control science and engineering from Zhejiang University, Hang Zhou, China, in 2004 and 2009, respectively. He is currently a Professor with the College of Control Science and Engineering, Zhejiang University. His research interests include networked sensing and control, cyber-physical systems, and control system security.

Jiming Chen (Fellow, IEEE) received the B.Sc. and Ph.D. degrees in control science and engineering from Zhejiang University, Hangzhou, China, in 2000 and 2005, respectively. He is currently a Professor with the College of Control Science and Engineering, Zhejiang University. His research interests include the Internet of Things, sensor networks, networked control, and control system security.